R E P O R T R E S U M E S

ED 003 899 48

PROBLEMS OF TURKISH LEXICOGRAPHY.

BY- TIETZE, ANDREAS
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER NDEA-VI-16-10 PUB DATE 60
EDRS PRICE MF-$0.09 HC-$0.64 16P.

DESCRIPTORS- *TURKISH, *LEXICOGRAPHY, *HISTORICAL REVIEWS, *PROBLEMS, *LANGUAGE RESEARCH, BLOOMINGTON, INDIANA

THE LEXICOGRAPHICAL PROBLEMS IN THE TURKISH LANGUAGE WERE DISCUSSED. HISTORICAL REVIEW OF THE LANGUAGE WAS PRESENTED WITH PROBLEMS OF LEXICOGRAPHY THAT EXISTED IN THE PAST COMPARED WITH THOSE OF THE PRESENT. DISCUSSION TOPICS OF THE REPORT INCLUDED (1) NAME OF THE LANGUAGE, (2) DELIMITATION COMPARED WITH RELATED LANGUAGES, (3) DELIMINATION COMPARED WITH EARLIER HISTORICAL PHASES, (4) DELIMITATION TOWARD NEW DEVELOPMENTS, (5) BORROWINGS FROM OTHER LANGUAGES, AND (6) DELIMITATION COMPARED WITH SUBSTANDARD WORD FORMS, SLANG, AND COLLOQUIALISMS. SPELLING, SCRIPT, AND COVERAGE PROBLEMS WERE ALSO REVIEWED. THIS REPORT IS ONE OF A SERIES OF 13 PAPERS PRESENTED AT A CONFERENCE ON LEXICOGRAPHY, INDIANA UNIVERSITY, NOVEMBER 11-12, 1960. (GC)

ED003899

16-18

DEPARTMENT OF HEALTH, EDUCATION AND WELF
Office of Educa
Document has been reproduced ex... ...ved fro
...n or organization originating it. Po... ...v or opi
...ted do not necessarily represent offici... Office of Educ
position or policy.

PROBLEMS OF TURKISH LEXICOGRAPHY

The lexicographer of any language, constantly forced to compromise between theoretical stipulations and practical considerations, will always have his problems. But there are languages where he has special problems, and one of them is Turkish. I shall summarize the problems of the Turkish lexicographer under two headings: (1) Problems of the language itself. (2) Problems arising in lexicography.

1. Problems of the language itself.

Turkish is a language with problems. These problems are essentially of an extralinguistic origin: they have to do with historical, cultural, and social phenomena, by which one's concept of the Turkish language is shaped, but they also effect the language itself deeply. Before the lexicographer can sit down to start his work, he has to make many preliminary decisions, and in these decisions he will follow his own concept of the language he is gogng to treat. Thus, the problems of the language will become the problems of the lexicographer.

Turkish is the language of the Turks in the narrower sense, of the people of Turkey. This sounds much simpler than it is. Linguistically, Turkish belongs to the large group of Turkic languages, which is itself a branch of the Altaic branch of languages. Culturally, the Turks have been an integral part of the Islamic culture area for the last 1000 years, and only within the last century has their attitude toward this cultural affiliation become somewhat ambivalent. Socially, the Turkish of the last 500 years was, in the first place, the language of the Ottoman dynasty (therefore called Ottoman), of the court, of the bureaucracy, of

the capital. In harmony with the classic cultural tradition of Islam, the art of literary expression was cultivated from the beginning and there was always a deep rift between the spoken language and the lofty style of writing, in which Persian and Arabic vocabulary become dominant (the borrowings of morphological and syntactical elements were less conspicuous). The extensive absorption of linguistically alien elements (Persian is an Indo-European, Arabic a Semitic language) led to structural changes in Turkish itself. Turkish became a language with two phonemic systems, one for words of Turkish origin, the other one for loan-words.

With the political changes of the last 100 years, the concept of the language changed also. First came the break with the tradition of the high literary style: with the development of journalism and the appearance of new political groups with an urge to spread their ideas to the masses, the old esoteric style of writing became outdated and a kind of standard colloquial was more and more used in its stead. Since the beginning of this century this development was already partly tinged with nationalistic objectives: the Ottoman language should be purged of the foreign elements and re-Turkicised. World War I and its consequence, the dismemberment of the Ottoman Empire led to a revolution of the Turkish way of life. The Kemalists of the 1920s and 1930s vigorously rejected the past, the cultural tradition of Islam, the dominance of the capital, the highbrow esthetic concept. They rejected the old language, which was the result of all these objectional factors. In 1928, by introducing a Latin alphabet and outlawing the Arabic alphabet, they forcibly and irrevocably cut off the old cultural tradition

and all ties with the Islamic world. They also announced a radical reform of the language itself. Since 1935, many long lists and entire books full of proposed new vocables for all the walks of life were published. In the late 'thirties, daily papers would publish their articles of comment on the political situation with the new vocabulary. The commentator would write his article in 'conventional' language, then a clerk would process it, replacing all Arabic (and Persian) by their 'Turkish' equivalents proposed in the official lists. However, he would add most of the eliminated words in parentheses after the new word, or in footnotes. As many of the new words were completely incomprehensible without annotation, the reader of the article had to rely strongly on the notes, or on a copy of the official wordlist (this time the one from Turkish to Ottoman). Undergoing this exercise every day, he would eventually learn to achieve a part of this manipulation in his mind: his eyes would read one word, but, mechanically, his mind would register another word, a mental habit which might be termed schizoglossia. In spite of the enormous energy that went into this matter, and the powerful official backing, only a part of the proposed vocabulary actually caught on. The words that entered the general language were relatively few, the great bulk pertained to special terminologies, mainly the terminologies of public administration, of the army, and of the sciences which are taught in high schools and colleges. The scientific terminologies had been very artificial before, too, when their elements were mostly borrowed from Arabic. Now, after the change of the alphabet and the new orientation of the whole system of education away from Arabic and Persian and toward the Western world, a continuation of

the earlier systems of terminology had become impossible. The result was a complete rift between the older and the younger generations of doctors, lawyers, engineers, etc.; the younger colleague had no way of discussing his problems with the older colleague. The younger generation started as a minority, but now the rift is in the mid-forties and they have already started to outnumber the older generation, who, after two more decades, will practically have died out. Thus, the Turkish language, both as a system, as a set of rules for communication, and as a social institution, with the relative distribution of the areas of application of these rules, is at present in a stage of transition. However, this evolution is not so straight and steady as one might conclude. A continual tampering with the language, although it was a historical necessity, has created contrasting attitudes of reaction toward stability and change: on the one hand receptiveness for new fashions, imaginative creativeness, a bold rejection of past achievements, especially along the line of esthetic refinement, on the other hand, a sneering rejection of everything that is new, an emphasis on classicism and traditionalism. The constant fluctuations also produced a feeling of instability, which, cut off from traditionalism, sought for compensations in the language of the people, in regional dialects, in slang, and, in the higher spheres of thought, heavily relied on French, German, or English. A certain degree of impoverishment and of Vergröberung had to be taken for granted by everyone.

This state of the language, with its conflicting concepts, creates serious problems for the lexicographer. Which of these concepts is he going to follow? Can he combine more than one

concept or all concepts? Before he begins his work, he has to make a number of decisions concerning the delimitations of his task. In each of these decisions, the problems inherent in the present stage of language with its conflicting concepts will stand in his way.

a) Name of the Language: The Turks call their language Turkish, but they speak of Ottoman when they specifically refer to the language of the older generation or to the older literature. An approximative borderline between the two might be the end of the Ottoman Empire and the foundation of the Republic of Turkey (1923) or the abolition of the Arabic and introduction of the Latin alphabet (1928). But the term Turkish, of course, was used long before the foundation of the Turkish Republic, even long before the foundation of the Ottoman Empire, for that matter. In fact, it was known in the West (in Byzantine chronicles) since the 6th century AD. There is enough difference between the Turkish of the rock inscriptions of the 8th century AD. and the language of Turkey of today, to necessitate a distinction in terms. Even scholarly terminology is not precise in this respect: We use the somewhat artificial term Turkic when we refer to the entire group of languages, including, e.g., Tatar, Bashkir, Kirgiz, Yakut, Uigur, etc. But this term is hardly used outside of linguistics: the historian speaks of the Turkish empire of Central Asia, the anthropologist speaks of Turkish shamanistic beliefs, etc. Not even the linguist is consistent, quite often he calls the language of the 8th century inscriptions Old-Turkish instead of Old-Turkic. For the language of the Ottoman Empire and of Turkey, the scholars use various terms: Ottoman, Ottoman-Turkish, Osmanli, and even

Osman, furthermore Turkish, Turkey-Turkish, Empire-Turkish (Reichstürkisch), and, in French, le Turquien (on the model of le Parisien or l'Alsacien). When a child has so many names, there must be something wrong with him. Moreover, the child himself does not accept any of these names, with the exception of the ambiguous Turkish and, perhaps, Turkey-Turkish (or Turkish of Turkey), which is innocuous, if unwieldy as a term. As, from the linguistic point of view, there is no justification in using one term for the language up to 1923 or 1927 and another term for the current stage of the same language, the simplest and most convenient way would be to continue using the term Ottoman, but here the Turks who have made the new concept of the language their own balk. Thus, already in naming the language, the lexicographer has reached a deadlock and has to compromise.

b) Delimitation vis-a-vis related languages: For some of the lexicographers of the second half of the 19th century, who in a romantic way delighted in the sense of unity of all the Turkic languages, appealing vocables or expressions of Jagatai and other non-Ottoman Turkic languages had a right to enter the (otherwise Ottoman) dictionary, and this highly arbitrary selection of intruders was again and again taken over by later lexicographers. As dictionaries usually stand firmly on the shoulders of older dictionaries, it takes a good deal of effort to discard materials which are so well attested by all the earlier standard works. Today, nobody would think of including Jagatai or Tatar words in a Turkey-Turkish dictionary. Not even other languages of the Southern Turkic or Oghuz subgroup, as Azerbaijani or Türkmen, would be considered. The only point where delimitation on linguistic

grounds is difficult is with the Gagauz dialect of northern Bulgaria, the Dobrugea, and Moldavia. But the practical importance of this problem is negligeable.

c) Delimitation vis-a-vis earlier historical phases: How far back ought the lexicographer to go in covering the language? For Turkish, a range of widely discrepant answers can be given, depending on the concept of the language: on the one hand, a word used by the popular Anatolian folk poet Yunus Emre around 1300 AD. can be regarded fit to enter, on the other hand, a word commonly used as recently as the 1930s, but stigmatized by the language reform movement, may be rejected as obsolete. The vocabulary used by Ataturk in his speeches is largely incomprehensible for the younger generation of today; a dictionary which contains it would already be regarded as a historical dictionary. The lexicographer who includes 'old' material (from 1300 to 1935) will have to be very careful to provide each word with an appropriate label: learned, obsolete, archaic, only known to the older generation, etc.

d) Delimitation toward new developments: The problem what to do with newly coined words or expressions will be equally difficult. Has anything proposed by anyone to be listed? Or, if the criterion is the degree of reception and assimilation by the language, how can this be measured? How can the schizoglottical pitfall be avoided? It will be very difficult to set up guiding principles of an objective character, and here again the lexicographer will have to rely on his discretion. No wonder that the policy of recent Turkish dictionaries in this respect varies radically. Some exclude the new material altogether, others link it by crossreferences firmly to the now rejected equivalents, thus

introducing the principle of schizoglossia into the dictionary, others, again, apply various principles of selection.

e) Borrowings from other languages: For many centuries, and to a large degree even today, Latin (and, to a lesser extent, Greek) has been the Kultursprache of the European languages, a lexical reserve arsenal from which anybody at any time could borrow the tools for expressing a new idea. For Turkish, both Arabic and, to a lesser extent, Persian have functioned as Kultursprachen for many centuries. An Ottoman writer or poet had the liberty to borrow from these languages any noun or adjective he needed (the borrowing of other word classes was more restricted). Consequently, if an Ottoman dictionary-maker wanted his dictionary to be adequate, he had to include, apart from the Turkish words and from the loanwords used in Turkish, all nouns and adjectives that he might find in an Arabic or Persian dictionary. This is the reason why Meninski called his Turkish dictionary, published in 1680, Thesaurus linguarum orientalium, turcicae, arabicae, persicae, or Handjéri, in 1840-41, Dictionnaire français-arabe-persan et turc. Even James Redhouse (A Turkish-English Lexicon, publ.1890) could enter the entire noun and adjective material of the Arabic Kamus and of the Persian Burhan-i Kati without having to bother to find out which of these words, or which of their various meanings, had ever actually been used in Turkish; he was justified to do this, because all this material was, potentially, at least, a part of the Ottoman language. Today, after the reorientation of Turkish culture, the potential aspect of this material has vanished, and the lexicographer has to approach it from a historical point of view: his guiding principle has to be to find out what has

actually been used and what not, and the latter part he has to discard altogether. For this the first requirement is a comprehensive Belegwörterbuch. Unfortunately, no such thing exists. All the attempts made so far were concentrated on the Turkish materials. There is no hope that we will have such a work for many years to come; therefore, in this field, again, the lexicographer will have to go by his more or less arbitrary second sense, and the outcome will be a reflection of conflicting concepts of the language.

With the reorientation of Turkish culture, Arabic and Persian dispppeared as Kultursprachen, but, though in a lesser degree, the Western languages stepped in, primarily French, but partly also English, German, and Latin. Whenever a Turkish scholar or scientist, banker, businessman, engineer, or even sports expert, wants to make a very clear statement, he used a French term (even if he does not know French and if his training has been in another language). With the instability and confusion of Turkish terminology in many fields, this will happen quite often. Thus, French is, to some extent, the Kultursprache of modern Turkish; its abstract nouns, at least, can be regarded as potential Turkish vocabulary. How much of this is to be taken in the dictionary? Some lexicographers exclude all 'international' terms, presumably, because they are understood anyhow. Is this a correct principle? Also from a practical point of view it can have bad consequences: think of an American travelling through Turkey with a car and a Turkish dictionary, who has engine trouble; it will be little help to him to learn that all the terminology of the automobile is, anyhow, in French.

f) Delimitation vis-a-vis substandard forms, slang, and regional vernacular: As to substandard forms and slang, the problems of a Turkish lexicographer will be more or less the same as those of lexicographers of other languages. But in his attitude toward dialects, his conceptual conflict will, again, come strongly to the fore. One of the revolutionary changes in the concept of the Turkish language within the last 50 years is the rejection of the cultural (and linguistic) hegemony of the capital (Istanbul). For many centuries, Istanbul, as the city of the Court and of the central bureaucracy, had dictated to the country its standard language. In the provinces, only the representatives of the capital, administrators, army officers, teachers, doctors, etc., and their families, spoke the standard language (and were called Ottomans, Osmanli, by the local population). Today, with compulsory education, compulsory military service, newspapers, radio, and, in general, more moving about, the dominance of the Istanbul standard language is more deeply entrenched than it ever was. But, on the other hand, the political dethronement of Istanbul, the rejection, on ideological grounds, of the grand tradition it stood for, and the much closer and more intensive contacts between the cities and the countryside have resulted in a vigorous influx of regional vocabulary into the standard language, expecially in literature. Today, it has become impossible to read a contemporary novel with the help of any of the existing dictibnaries. Here, new tasks arise for the lexicographer, and, again, he is confronted with the problem how far he should go in covering this new material of regional proveniency. An almost infinite number of attitudes are possible: he could go by the frequency of

occurence in literature, using some statistical yardstick; he could go by the meaning, e.g., only accepting terms for things or ideas of local character; or he could divide the material by categories, accepting, e.g., proverbs, sayings, idioms, etc., perhaps also terms, but rejecting merely phonetic variants, etc., etc.

2. Problems arising in lexicography.

As we have shown, the problems resulting from the specific situation in which the Turkish language finds itself today are at the same time problems of the lexicographer. But there are also problems more closely connected with lexicography itself. I shall only list a few of these.

a) Problems of coverage: Above, in listing the problems of a proper delimitation of the historiographer's task, we have already, indirectly, outlined the difficulties with which he has to cope. After having made his decisions as to how to delimitate the area he wants to cover, he will have to bring the materials within this area together. As his policy will probably not coincide with that of any of his predecessors, he will have to be very careful when using older dictionaries. He will have to discard thousands of items, meanings and entire entries, which have become obsolete or which may actually never have existed at all in Turkish, including the 'potential' material from the older Kultursprachen as well as neologisms, proposed but soon forgotten. On the other hand, he will have to cover new territory, for which he cannot rely on previous dictionaries nor even, sometimes, on any preparatory study. He will find the newspaper, fiction, the talk of people in the street, songs, all kind of folklore, etc., etc., inexhaustible sources. The tradition of the Turkish

dictionary is of the learned type; when it comes to the vocabulary of practical life, of the bazaars, of the various trades, sports, games, even of the machinery of public administration, little preliminary work has been done and the lexicographer has to be himself an explorer.

b) Problems of the script and spelling: The change, 32 years ago, from the Arabic to the Latin characters is another source of problems for the lexicographer. It is difficult to discuss these problems without becoming too technical. May I only give a few hints. The new (Latin) alphabet is not just a transliteration of the old (Arabic) with differently shaped symbols. The relation is much more complicated. For instance, one letter of the Arabic alphabet may correspond to four letters of the Latin alphabet (as in the case of k, g, n, and ğ, the latter, itself, having two different pronunciations), but also one letter of the new alphabet (e.g., z) may correspond to as many as four letters in the old alphabet. Moreover, vowels were often not indicated at all in the Arabic characters. Consequently, it is not possible from the mere sight of a word in Arabic script to know how its spelling would be in the Latin characters, or vice versa. It is therefore practically impossible to use a dictionary in the modern Latin script for a text in Arabic script, and vice versa. The only possibility of producing a dictionary that could be used for both scripts is by giving two complete alphabets, one Arabic, one Latin, of which one could be used for the discussion of the entries and the other for crossreferences (the way it was done, e.g., in Bertels' Tajik-Russian dictionary of 1954, or in Borovkov's Uzbek-Russian dictionary of 1959).

One other consequence of the change of the alphabet is the instability and fluctuation of spelling habits and spelling rules. Although these concern mostly points of minor importance, they are a serious problem for the dictionary-maker. Should he just stick to one set of rules which happens to be accepted by the majority at the time when he starts his work, and ignore all deviating variants, or should he list them, too, with crossreferences?

A special headache are foreign words and foreign names. The new Turkish alphabet, though based on Latin, has some extra letters (e.g., ş for sh, ç for ch, etc.) and does not permit combined letters for the representation of one sound (like sh, ph, etc.), it does not have the letters x and w, and y occurs only as a consonant. For this reason, Turkish is not so conservative in the spelling of foreign words as English; like Italian and other languages it tends rather to reproduce their sound picture(French guichet is spelled gişe, duy is the French douille, etc.). Proper nouns have a difficult fight against this tendency (Avusturya for Austria is standard, but one sees Şikago besides Chicago, Kan besides Cannes). How are famous persons of European history to be called and to be spelled? Is it more correct to call the great emperor Charles V or 5'inci Şarl or (as it is often done) Şarlken?

c) Problems of definition andequivalent: A number of problems arise when the lexical elements of one language have to be defined in terms of another language of entirely different character, e.g., when the elements of Turkish, an Altaic language, have to be defined in English, an Indo-European language. Let me give a few examples: The Turkish noun has a marked form only for the plural,

the singular is an unmarked form which can refer to a single thing, but in many cases also to a plurality. Thus adam means both 'man' and 'men, people' at the same time. Furthermore, there is no definite article, the noun has a marked form only for indefiniteness. Thus, adam means 'man' or 'the man' or even 'a man'. This shows that already for syntactical reasons Turkish adam (or any other noun) has no exact equivalent in English. The Turkish adjective has lost its special form for the comparative. Therefore, Turkish çok means 'many', but at the same time also 'more', and even 'too many'. Turkish, also, does not distinguish between adjectives and adverbs, thus çok can also mean 'much' and 'very'. Nor does Turkish distinguish between adjective and noun, so that çok is also 'the greater part, most'. Again, there can be no exact equivalent in English for any Turkish adjective. I could go on with the other word-classes, but these examples may suffice.

Turkish, as an agglutinative language, has at least one advantage over the Indo-European languages: it considerably facilitates the task of integrating the lexical dictionary into a dictionary of morphemes. If we disregard a negligiable number of exceptional cases, it has no intramorphemic changes and intermorphemic juncture problems are minimal. Furthermore, morphemes have, in general, retained their free affixibility and not lost their meaning. The number of their possible combinations is very high. In a morpheme dictionary, even of limited volume, the student would find the tools with the help of which he could analyse the most complex derivative or grammatical form, and even, if he is lucky, derive himself more or less correctly the word form he needs.

Andreas Tietze
University of California
Los Angeles, California